DOCUMENT RESUME

ED 024 917 AL 001 253

By- Crowley, Dale P.

The Occurrence and Function of Chinese Characters in Modern Japanese Orthography.

Pub Date 27 Apr 68

Note- 15p.; Paper presented at the 21st University of Kentucky Foreign Language Conference, Lexington, Kentucky, April 27, 1968.

EDRS Price MF-$0.25 HC-$0.85

Descriptors- *Chinese, *Japanese, Morphemes, *Orthographic Symbols, Phonemes, *Writing

Identifiers- Kanji, *Morphographs, Phonographs

One of the objectives of this paper is to point out that the essential function of Chinese characters for the reader of Japanese is that of mediating, or evoking the sounds of Japanese directly, and the meanings secondarily, as the meanings are derived from the sounds. The author suggests that this is what happens, despite the fact that Chinese characters are known as "ideographs" (graphic representations of objects or thoughts); he prefers to call Chinese characters "morphographs"--they elicit auditory responses, or phoneme sequences, and function as graphic representations of the morphemes of the language. From these phoneme sequences, whether verbalized or silently conceptualized, come the meanings or ideas being communicated. The ultimate step in the evolution of graphic symbols from representations of objects and thoughts to representations of morphemes is the complete cancellation of any connection that the graphic symbol has with meaning (as in the English alphabet and the syllabary systems of Japanese). Topics dealt with are the Kanji used by Japanese scholars as Chinese words, the use of Chinese characters as phonetic symbols, the introduction of Chinese loan-words, homophony, modern occurrence and function, and recent investigations. In summary, the author states that the occurrence of approximately 2,000 Chinese characters in Japanese writing is excessive and inefficient. (AMM)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Dale P. Crowley
University of Hawaii
Peace Corps Training Center
Hilo, Hawaii

ED024917

THE OCCURRENCE AND FUNCTION OF CHINESE CHARACTERS IN MODERN JAPANESE ORTHOGRAPHY*

Dale P. Crowley
University of Hawaii

Interest in Japanese orthography might stem from a belief that literacy is important for the learner of the language, from an interest in possible pedagogical procedures, from questions about the discrepancies between spoken and written usage, and from curiosity on frequency of occurrence of Kanji in Japanese.

One of the objectives of this paper is the attempt to point out a thesis related to the topic--namely, that the essential function of Chinese characters for the reader of Japanese is that of mediating, or evoking the sounds of Japanese directly, and the meanings secondarily, as the meanings are derived from the sounds. I will take this opportunity to draw attention

---

*This paper was presented at the 21st University of Kentucky Foreign Language Conference at Lexington, Kentucky, for the program "East Asian Languages," on April 27, 1968.

AL 001 253

to this theory since it bears directly on the function of the characters. Considerable attention has been focused on these phenomena in the reading of languages with simple orthographies, such as English, but I suppose that perhaps because of the complexity of Japanese orthography most scholars have concluded that reading Japanese is different from reading English and other simply written languages.

To state the theory clearly, the major, or essential function of Chinese characters in Japanese writing (or, in Chinese writing for that matter) is that of evoking the phoneme sequences of the language for the reader. From these phoneme sequences--whether verbalized or silently conceptualized--come the meanings or ideas being communicated.

This is what actually happens, I suggest, despite the fact that the characters are best known as ideographs. The definition of ideograph is sufficient to convince us that Chinese characters were at some time in history (just as the letters of our alphabet were) graphic representations of objects or thoughts.

A corollary of the thesis is that true ideographs cannot remain so for long; that is, so long as they are used in human communication. We must first assume, in order to support this corollary, that the objects and thoughts for which graphic representations have been invented are represented also by phoneme sequences in oral language. This relationship between two stimuli--one visual and one auditory, and one response--the morpheme of oral language, sets up a semblance of the classic environment for the development of conditioned responses.

That is, the graphic-visual stimulus, originally acting as an evoker of an idea response would, in human communication, inevitably occur simultaneously with the auditory stimulus. Since the auditory stimulus also

evokes an idea response, both stimuli are tied to the same response. Eventually the graphic-visual stimulus acquires the potentiality of evoking not only the idea response, but the auditory stimulus as a response also. The graphic-visual stimulus thus begins to evoke the auditory response, which in turn, as a stimulus, evokes the idea response. (Many scholars have discussed this phenomenon as it occurs in the reading of English.)[1]

For this reason I prefer to call Chinese characters morphographs. As elicitors of auditory responses, or phoneme sequences (single phonemes in a very few cases), they function as graphic representations of the morphemes of the language. I will concede that an idea or logos may exist apart from the phonemes of language, but morphemes cannot. Consequently, according to my theory, since Chinese characters become so directly and strongly tied to phoneme sequences and morphemes, they are more properly labeled morphographs. (In this paper morphograph is synonymous with Chinese character and Kanji.)

The ultimate step in the evolution of graphic symbols from representations of objects and thoughts to representations of morphemes is the complete cancellation of any connection that the graphic symbol had with meaning. Such is the case with our own alphabet, as well as the syllabary systems of Japanese. I am convinced that many Chinese characters appearing in Japanese orthography no longer bear any meaning for the reader, but serve only as elicitors of a phoneme sequence.

Of course, we must allow for the tendency that we have--perhaps the conditioning referred to earlier working in the other direction--to see written words as the embodiment of the idea they ultimately represent. The word "castle" just looks nicer than the word "shack." But I feel that this kind of artistic or esthetic consideration is paralinguistic, and can be discounted for the purposes of this discussion. Nevertheless, it is easy to

see certain characters as the very epitome of the morpheme they represent, the propitious appearance of the Kanji kotobuki, the happy shape of the Kanji tanoshimu, and the wicked profile of the Kanji tsumi, for example.

Kanji Used by Japanese Scholars as Chinese Words

We will now review the history of the occurrence of Chinese characters in Japanese writing, and tie the facts in with modern usage and function where significant.

The first contact that Japanese speakers had with Kanji was that of bilingual scholars handling the lexicon of Chinese, around 200 A.D., give or take a hundred years.[2] This is parallel to the modern scholar who utilizes the lexicon of a foreign language as it is, without any attempt to incorporate it into his native language.

Within a hundred years, however, these scholars were experimenting with Kanji as carriers of meaning in a Japanese sentence, retaining the Japanese morphemes. The characters were written in the Japanese grammatical order, with the particles and inflections also represented--very clumsily--by Kanji, which in this case had little or no semantic connection with the structural meanings of Japanese grammar. Of course, only the scholars who knew the meaning of the characters could read the writing, but the sound patterns of Japanese were not altered in the least. There is a modern counterpart, to be sure, in the use of Chinese characters as orthographical representations of native Japanese words. The most notable examples are family names, place names, and such common words as tomodachi, kimono, deguchi, etc.

The Use of Chinese Characters as Phonetic Symbols

The second, more linguistically significant development, was the use of Kanji for their sound value (and the word value is used advisedly, as there are no inherent sound values in any graphic symbol). Now this feat was not without enormous difficulties. In the first place, the Chinese phonemic system was radically different from Japanese. To account for this discrepancy the sound of the characters was modified to fit the Japanese phonemic structure.[3] What existed, in effect, was a system for using certain Chinese characters for Japanese syllables because those characters represented similar sound sequences in Chinese. It should be stressed, the characters were not used for their meaning.

Perhaps the greatest potential trouble with this system was that as many as a dozen characters that had pronunciations similar to the syllables of Japanese could be made to represent the same sounds. The learning load for the reader was enormous, but the characters were used for their phonetic, not semantic function. The sound of the Japanese language remained the same, consisting of basically the Consonant-Vowel sequence. With this, Japanese thus had the beginnings of a purely phonetic writing system.

It is worth noting for its pedagogical implications that through a process of conditioning through reading and rereading, the habit-strength tie between the Kanji and the phoneme sequences of Japanese rendered the Kanji phonographs. This, as I have already suggested, is precisely their function in modern Japanese. That the human mind can assimilate and systematize such a conglomeration of complex graphic shapes, and respond to them with auditory sequences(according to my theory) from which meaning is derived is phenomenal.

One of the happy outcomes of this process was the systematization of a favorite set of phonetic Kanji, now known as Man'yoogana, used with consistency in the Manyooshuu. As a matter of fact, several such sets of original standard representations of Japanese sounds developed through the centuries, some of which although little known, remain in limited use to this day.[4]

The significant phenomenon in all this is that the graphemes were the servants of the phonemes, not vice versa. If the characters being employed to represent sounds were too complex, elements were abstracted, or the whole character was written in a cursive shorthand. Thus were born the syllabaries. This further development of the syllabaries, which included the organization of the Man'yoogana was, to my mind, one of the most versatile and sophisticated, linguistically significant adaptations of writing in the history of language. Whether Kanji, or kana derived from Kanji, the graphemes, I suggest, serve the morphemes of the language.

In conclusion to this section, I consider this historical development of the evolution of sets of characters which represent the syllables of Japanese speech to be major support for my theory that the basic function of Chinese characters in Japanese orthography is still that of eliciting phoneme sequences for the reader.

The Introduction of Chinese Loan Words

The next major historical development was the introduction of Chinese loan words accompanied by their own characters. This is the development that, in my opinion, made Japanese orthography the formidable learning task that it is today. So long as Kanji were used for representing phoneme sequences only, their number was minimal, or at least easily controllable.

But when the practice of incorporating Chinese morphemes and their accompanying characters into the language caught hold, the number of Kanji appearing in Japanese writing proliferated to around 50,000 by the 19th century. (The need for and function of so many Chinese loan words, many alongside Japanese words of identical meaning, is an interesting subject in itself.)

Homophony

An interesting sidelight to this development was the literal multiplication of homophones. Traditional Japanese, or the Yamato Kotoba, had comparitively few homophonous morphemes. But with the introduction of Chinese loan words, tones and other phonemic entities that were significant in Chinese were lost in the process of phonemic restructuring for Japanese. The outcome of this was literally hundreds of sets of homophones, some sets consisting of as many as a score of words. Despite modern reforms which have reduced the number of morphographs drastically, I found that the approsimately 2000 General Use Kanji are used to represent a few more than 2000 (2011) morphemes and allomorphs. But these morphemes are encoded by only 286 different quasi-Chinese phoneme sequences, or a ratio of seven morphemes and allomorphs to one phoneme sequence. It is clear that the potential number of homophones possible from combinations of so many homophonous morphemes is astronomical. The phoneme sequence /koo/, for example, is represented by 60 morphographs in modern Japanese orthography, /shoo/ by 59, /kan/ by 43, and so on. Any combination of these three syllables alone would produce a half-dozen or more words in current use.

The study of homophony (if morphographs are used to write the word), or homonymy (if a syllabary is used) in Japanese is fascinating. I

feel that the psycholinguistic implications of such a system deserve more attention. The facts that the average student--native or foreign--does not have more difficulty assigning the same phoneme sequence to so many different morphemes and morphographs, and that so few ambiguities occur in oral communication are, to my mind, startling.

Modern Occurrence and Function

Thoughtful reformers have been busy for over a century with Japanese orthography, and many of their recommendations have been accepted. The number of Kanji in popular use has been steadily reduced. The most recent cut was the result of legislative action in 1946, setting the number of Chinese characters for general use at 1850, or 1878, or, including those utilized for special needs, approximately 2000. 881 of this number were selected for systematic presentation during the first six years of school, the remainder considered to be essential for all ordinary reading tasks.

Now it is interesting, and bears on the subject, I think, to note that these decisions were made without the benefit of a computerized investigation of current usage at that time. Although I am certain that the major criterion for this selection was usage, I think we can ask ourselves, even at this late date, how accurately were actual written and oral usage determined? Certainly we need not worry about the inclusion of several hundred omnipresent morphographs, but I continue to wonder about several hundred at the low end of frequency of occurrence, that might better have been included, or excluded. My studies have shown that several hundred of the so-called general use characters are used to represent only one or two words in modern Japanese, and that these words occur for the most part with very low frequency.

Recent Investigations

Eleven years after the legislation that limited the use of Chinext characters, the National Language Research Institute, a government agency, published its investigation of the frequency of occurrence of lexicon in popular Japanese magazines.[5] The sample was 13,000,000 words.

Three years later, 14 years after the legislation, the National Language Research Institute published the results of its investigation of the frequency of occurrence of the morphographs in the same periodicals.[6] The sample was 116,000 morphographs. By correlating the information of the 1957 word count with the 1960 morphograph count, I discovered much significant information on the occurence and utilization of Chinese characters in modern Japanese. I have space to scarcely mention the more important data.

By combining the information of the two reports; that is, determining which high frequency words were represented by which high frequency morphographs, I found that the first 500 morphographs, or 25% of the approximately 2000 in general use, account for 75%, or 87,000 of all occurrences of all morphographs in the sample, and are used to represent some 2,500 of the most important words in modern Japanese. (For the sake of simplicity, most figures are rounded to the nearest 100, some to the nearest 1000. High frequency, with reference to lexicon, means at least one occurrence in 20,000.)

The implication of this data is that the use load is on a comparitively few characters. We can see this picture better by reaking the morphographs down into groups of fifty.

The first fifty most used morphographs occurred 30,000 times in the 116,000 sample, or 24% of the sample. These occurrences included the representations of over 500 of the most used words in Japanese writing.

The next fifty occurred less than half as often as the first fifty--approximately 15,000 times for 12% of the total sample. These morphographs were used to represent over 300 high frequency words.

At this point we have 100 morphographs, or 5% of all those in general use accounting for 36% of all occurrences.

The third group of 50 occurred almost 11,000 times in the 116,000 sample for 9%, and were used to write about 250 high frequency words.

The fourth group of 50 occurred over 8,000 times for 7%, and represented over 200 high frequency words.

At this point we have 200 morphographs, or 10% of those in general use accounting for 52% of all occurrences. I will not continue with this detailed data, but the evidence is dramatic that a few hundred morphographs are very important in Japanese orthography.

This is not to say that the vocabulary of 2,500 high frequency words represented by the highest ranking 500 morphographs is sufficient, or that the remaining 1,500 morphographs need not be studied. I will venture to say, however, that while the use of the first 500 does indeed feature many advantages to the speaker and reader of Japanese, after that, whether the advantages gained are commensurate with the effort that must be invested in learning is indeed questionable, particularly when (1) the hazards of homophony diminish at the lower end of the rankings, and (2) the syllabary can be used to write anything in Japanese.

To give you some idea from my data of the kinds of situations that obtain with particular characters, I mention these typical cases briefly:

- Morphographs that are important, but restricted in their load of lexicon representation are such as the SAI of KEIZAI, 131st rank, with nearly 200 occurrences in the word KEIZAI, and none in any other high frequency word.

The SHI of SHIGOTO is 271st in frequency of occurrence ranking, occurring 99 times in SHIGOTO, and in no other important word. (It is recognized that this morphograph may be used to represent many words more common in spoken usage, such as shiage, and shiire.)

BOO in BOOEKI is 546th in rank, occurring 48 times in the sample, and all 48 in the important word BOOEKI.

My favorite in this category are both the KEN and the KYUU of KENKYUU. These morphographs are 284th and 261st in rank respectively, and occur in only one high frequency word, KENKYUU.

The point in these examples is that when the student learns certain morphographs, he is learning them for use in only one major word in the language.

Another interesting category is the list of very infrequently occurring morphographs, likewise extremely limited in the words in which they commonly appear. The SO of KISO occurred very infrequently, but when it did, usually in the word KISO. The TEI of TEIKOKU in that common word only, the SOTSU of SOTSUGYOO in that common word only, and the MI of MIRAI in that common word only.

An example of the influence of current thought on the occurrence of morphographs is found, I feel sure, in the SEN of SENSOO, 32nd in rank, occurring 473 times in such words as SENSOO, SENZEN, SENGO, TAISEN, HAISEN, and SHUUSEN.

Throughout all of this statistical investigation, which has provided much information and diversion, I must state that my primary interest has been in its possible pedagogical application. It seems to me that the studies of the National Language Research Institute, and my data based on it, have tremendous application in the development of language materials and courses.

Educational psychology tells us that immediate usefulness, ample reward reinforcement, short success steps, and immediate feedback are crucial to efficient learning. Yet, most courses in reading, with all their admitted relevancy, include some irrelevant, virtually useless morphographs as well as lexicon. I suggest that a core of most frequently occurring morphographs and lexicon, with interrelationships between the two, be identified, and lessons built around that core for presentation to all beginning readers of Japanese.

Educational psychology tells us further that for optimum learning the mature student is entitled to know, to the greatest possible degree, the rationale behind and the ultimate usefulness of what he is being required to learn. I propose that the rank order of morphographs, and the important lexicon, as well as the frequency of occurrence of the lexicon which they are used to represent, be indicated to the student of Japanese orthography. The result of these strategies can only be stronger goal-oriented study, and more efficient learning.

Furthermore, it is obvious from the data cited that the division between Kyooiku Kanji, and those that remain of Tooyoo Kanji, reflects many arbitrary judgments not based on actual importance and occurrence. And in my examination of reading texts for Japanese I have come to the conclusion that likewise, minimal attention has been given to objective studies of frequency of occurrence of either the lexicon of Japanese, or of the morphographs used to represent the lexicon. Selection has clearly been intuitive, not without validity, for our subjective "feel" for heavily utilized lexicon and graphemes is usually accurate.

I will never forget struggling with writing KABIN within a month after arriving in Japan, simply because it was included in one of the first

lessons in a certain course. But I rarely saw the word in print thereafter, only in that book. This represents a grossly wasteful application of the intellect, and should serve to enlighten us on much of the reason for the lament over the difficulty of Japanese writing.

Another Corollary of the Thesis

To refer once more to the proposition that we have considered throughout this paper, I would like to comment on the consequences of eliminating all morphographs from Japanese orthography. Not that I am in favor of such a measure. I simply want to consider some of the results forthrightly.

Another corollary of the thesis that evoking sound responses is the primary function of morphographs is that no serious damage would be done to communication among speakers and readers of Japanese if the use of all morphographs were abandoned. If the morphographs were direct evokers of meaning such would not be the case--their use would be essential to communication. But assuming the validity of the thesis, and barring some ambiguity arising from homophony, efficient communication could be maintained with any form of so-called phonetic writing. The ease of difficulty of any system of orthography is purely a relative consideration, the intricacies of the difficulty being mitigated by sufficient encounter with the system.

But this brings to mind another obvious, and highly desirable function of morphographs--the parsing of sentences into their grammatical components. Native speakers and foreigners alike experience the facility afforded to the task of reading by an orthographical system which uses morphographs to represent nouns and verbs and other major content categories, and the syllabary to represent minor lexical categories, inflections, and particles. Parting with this convenience as a consequence of abolish-

ing morphographic writing is one that many could not happily accept. Yet we must honestly concede the inevitable corollary of the premise--that since the morphographs are primarily and directly evokers of sound responses, the human intellect would soon accustom itself to another system of orthography, such as a purely syllabic or alphabetical one.

Summary

In summary, I will say that the occurrence of approximately 2000 Chinese characters in Japanese writing is more than enough to insure good communication. The advantages of using several hundred high frequency morphographs to represent several thousand high frequency words, dividing the written sentence into easy-to-read sections, and serving as a protection against ambiguity which arises from homophony, are readily recognized.

Furthermore, since homophony is more likely to occur among these 500 most used morphographs (because of the statistical weight on their use), little crucial ambiguity would result if the use of infrequently occurring morphographs were abandoned. This notion is further supported by the thesis that the Kanji used in Japanese writing actually represent Japanese speech, and as such evoke the sounds of the language primarily. Although we recognize the ability of the intellect to acquire responses for a virtually unlimited number of stimuli, we are forced to concede that learning to read and write two or three complex shapes, for example, in order to represent one, and only one word of common use in the language, when those shapes have no direct connection with meaning, is a most inefficient system.

Finally, any course in Japanese reading should be developed around these considerations of occurrence and function if we are to set up the best conditions for learning.

BIBLIOGRAPHY

[1]Leonard Bloomfield, Language (New York: Henry Holt and Company, 1950), p. 285.

Leonard Bloomfield and Clarence L. Barnhart, Let's Read (Detroit: Wayne State University Press, 1961), pp. 31, 36.

Charles C. Fries, Linguistics and Reading (New York: Holt, Rinehart and Winston, Inc., 1963), p. 205.

Donald D: Durrell, Improving Reading Instruction (New York: Harcourt, Brace and World, Inc., 1956), p. 159.

[2]George Sansom, An Historical Grammar of Japanese (London: Oxford University Press, 1928), pp. 1-4.

[3]橋本進吉, 国語概論 (東京:岩波書店, 1946), pp. 121-123.

[4]Ibid., pp. 124-130.

[5]The National Language Research Institute, Research on Vocabulary in Cultural Reviews, Part 1--Vocabulary Tables (Tokyo: The National Language Research Institute, 1957).

[6]The National Language Research Institute, The Use of Written Forms in Japanese Cultural Reviews (Tokyo: The National Language Research Institute, 1960).

(Instead of "Cultural Reviews," a more accurate translation from the Japanese titles of the above works is "General Magazines.")